Central Bank Stress Tests: Mad, Bad, and Dangerous

Kevin Dowd

> In my youth it was said that what was too silly to be said may be sung. In modern economics it is put into mathematics.
>
> —Ronald Coase

One of the most important aspects of the remarkable transformation of central banking following the onset of the 2008 global financial crisis is the growth of regulatory stress tests for the larger banks. The relevant regulator—typically the central bank—uses these to determine the ability of the banks to withstand stress, and uses the results of the tests to assess the overall financial health of the banking system. A key purpose of the stress tests is to reassure the public that the banking system is sound.

When putting banks to such a test, the relevant authority starts by imagining some stress scenario(s) to which banks might be exposed—these are effectively just guesses pulled from thin air—and uses a bunch of models based on a bunch of further guesses to determine how the scenario(s) will affect the banks' capital adequacy (i.e., their ratio of capital to assets) over the course of the

---

Cato Journal, Vol. 35, No. 3 (Fall 2015).

Kevin Dowd is Professor of Finance and Economics at Durham University, an Adjunct Scholar at the Cato Institute, and a partner at Cobden Partners in London. This article is based on his talk at "The Cato Summit on Financial Regulation" in New York City, June 2, 2015, and also draws upon material in Dowd (2014, 2015).

stress period. It then passes or fails individual banks according to whether their capital ratio has remained above some minimum by the end of that period. To take a typical example, in its latest (2014) stress tests, the European Central Bank (ECB) assumed a single scenario, took the capital ratio to be the ratio of Common Equity Tier 1 (CET1) capital to risk-weighted assets (RWA), and selected a minimum required ratio of 5.5 percent measured in terms of the CET1/RWA ratio. Any bank that maintained a CET1/RWA ratio of at least 5.5 percent by the end of the stress period was then deemed to have passed, and any bank whose capital ratio fell below this minimum was deemed to have failed.

These regulatory stress tests are the ultimate in the appliance of financial "rocket science" to the banking system, and many of the models themselves are derived from the physical science models used so successfully in real rocket science. However, by their very nature, all these models—the financial models and the stress tests themselves—are impenetrable black boxes to any outsider, and we are asked to take their reliability on trust. The analogy with rocket science, though appealing and even comforting, then breaks down in two critical respects:

- Real rocket science is grounded in the science of physics, and the laws of physics are well established. By contrast, so-called financial rocket science is merely a set of beliefs and practices based on sets of convenient assumptions that ape some of the assumptions made in physics, but are wide of the mark as descriptions of how financial markets really work.
- We know that the methodology underpinning real rocket science actually works because it is scientifically tested, but we have no such assurance with its financial and central bank equivalents. Indeed, going further, we can say, with confidence, that we know that the methodologies underpinning both financial models and regulatory stress tests do not work: the stress tests provide an extremely unreliable radar system.

My purpose in this article is to spell out this latter claim—or, more precisely, to assess the methodology of regulatory stress testing both by reference to first principles and by reference to its track record. The results are shocking.

Financial Risk Models Are Worse than Useless

The first point to appreciate is that central bank stress tests are based on models of financial risk—models that predict potential losses and their associated probabilities—and these models are not so much useless as worse than useless. More precisely, the stress tests are dependent on risk models because they make use of risk-weighted asset measures that are dependent on the risk models. These models are useless at predicting financial losses and worse than useless as risk management tools because of their gameability and the false risk comfort that they provide.

Consider the foundations of risk modeling. The first is the standard assumption that financial returns (or losses) follow a Gaussian (or normal) distribution. A nice example illustrates that this assumption is impossibly implausible for the large losses that really matter. Back in August 2007, Goldman Sachs' hedge funds were experiencing enormous losses. "We're experiencing 25-sigma [standard deviation] events, several days in a row," explained their CFO, David Viniars (Larsen 2007), the suggestion being that Goldman had been very unlucky as opposed, e.g, to merely being incompetent. Financial commentators were quick to pour scorn on this lame excuse, and 25-sigma events were being likened to events one would expect to see on one day in 10,000 or 100,000 years. That's a long waiting time for events that actually happen quite frequently in financial markets.

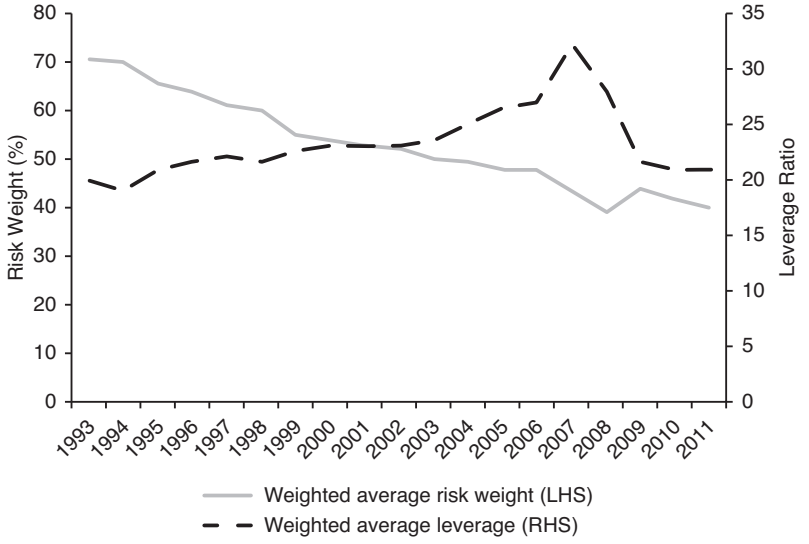
However, under the Gaussian distribution, the waiting time to observe a single-day 25-sigma event is much, much longer than even 100,000 years. In fact, the waiting time is 1.3e^135 years: 1.3 with 135 zeros inserted after the decimal point (Dowd et al. 2008: 3). To put this number into perspective, the number of particles in the known universe is believed to be somewhere in the region of 1e^80, which is literally infinitesimally smaller. To recycle an old Richard Feynman joke, a number like 1.3e^135 is so large that the term "cosmological" hardly suffices; perhaps we should describe it as "economical" instead. Thus, the Gaussian distribution massively underestimates the risks of the really big losses that truly matter.

A second pillar of risk modeling is the Value at Risk (or VaR) risk measure. This tells us the maximum likely loss that can occur on a position at a certain level of probability, for example, on 99 times out of 100. In plain English, this definition boils down to the worst we

can do if a bad event does not occur. Unfortunately, it tells us nothing about the loss we might experience if a bad event *does* occur—and it is the very high losses that we should worry about; the VaR is blind to the risks that matter, the ones that can wipe a bank out.

A third problem with the risk models is simply that they don't work. One could give many examples (see, e.g., Dowd 2014: 6–8), but Figure 1 suffices. The continuous dark plot shows banks' average risk weight, which includes the impact of risk models; the dashed line shows a primitive metric, bank leverage, the ratio of bank assets to capital, which ignores risk models. The risk-weight plot suggests that risks were continually falling; the leverage plot shows that they were rising up to 2008. As Bank of England economist Andrew Haldane (2011: 3) noted, "While the risk traffic lights were

FIGURE 1
Average Risk Weights and Leverage, Selected Banks, 1993–2011



Notes: The sample consists of Deutsche Bank, HSBC, BNP Paribas, Barclays, Citigroup, UBS, BAML, BONY, Commerzbank, ING, JPM, LBG, Santander, State Street, UniCredit, and Wells Fargo. Data are not available for the remaining G-SIBs. The leverage ratio is defined as total assets relative to Tier 1 capital.
Sources: Bank of England, *The Banker*, Bloomberg.

flashing bright red for leverage [as the crisis approached], for risk weights they were signalling ever-deeper green." The risk weights were a contrarian indicator for risk, indicating that risk was falling when it was, in fact, increasing sharply.

There are a host of reasons why the models failed so badly, but only one that matters: gaming. The models were being used not to manage risks, but to game the risk-weighting system. No model can take account of the ways in which it will be gamed, and market players have strong incentives to game the models used to control them.

So why does bad modeling persist? The reason is that banks want bad models because they understate their risks, and the regulatory system endorses bad models because it is captured by the banks.

Most risk modeling is then just a game: banks pretend to model risks, but they are really gaming the risk numbers. This game even has a name: risk-weight optimization. You fiddle with the models to get low risk numbers and you come up with clever securitizations to game the risk-weighting rules.[1] The lower the risk number, the lower the capital requirement, and the more capital can be siphoned off and distributed as dividends or bonuses.

In short, the real (though seldom publicly stated) purpose of risk modeling is to use capital regulations to decapitalize banks, and when they go bust, the bankers play dumb and lobby for a bailout. It is perhaps no wonder, then, that the risk models don't work: they are not intended to.

General Problems with Regulatory Stress Tests

Besides their dependence on risk models, stress tests carried out by central banks or financial regulators are subject to a number of other problems, any one of which on its own would be enough to render such exercises fatally flawed.

One problem is that these tests are essentially based on a single scenario.[2] Reliance on a single scenario violates good practice advice on stress testing and even common sense (see Finger 2008). Even if

[1]A good example is the "how to destroy" securitization. The purpose of this securitization was to game the Basel capital rules to release capital and generate false profits (Kerr 2010, 2011).

[2]Strictly speaking, the Fed in its Comprehensive Capital Analysis and Review (CCAR) stress tests uses three scenarios, but of these only its "extremely adverse" scenario really matters.

the banking system is safe under the one scenario that you considered, how do you know that it will be safe against all the other scenarios that you did not consider? You don't. The odds of that scenario actually occurring are also vanishingly small: whatever you think might happen, something else generally does. No single scenario can possibly give you confidence that the banking system is safe.

Another problem is that central bank stress tests are typically undemanding, that is, insufficiently stressful. They are undemanding in two different ways:

- The stress scenario itself is typically only a mildly adverse scenario.
- The hurdle (or pass rate) is (always) low, and arguably very low.

To give an analogy, a stress test is like an exam, and a tough exam requires not only a tough exam paper but also a challenging pass mark. Central bank stress tests are then like having an exam with an easy set of questions and a very low pass mark.

A third fatal error with central bank stress tests is that they lack credibility. They lack credibility in part because the regulators who conduct them are prone to capture by bankers, who then pressure regulators to go easy on them in much the same way that students would pressure a weak teacher to give them an easy exam, and in part because of central banks' own dismal forecasting records—none of them saw the financial crisis coming and they have all made major forecasting errors in the period since the onset of the crisis.

Stress tests also lack credibility for another reason. Imagine that a central bank conducted a stress test that suggested that the banking system was very weak. The central bank could never reveal such a result, because to do so would undermine public confidence in the banking system and violate one of its most important tasks, which is to maintain that confidence. Remember, too, that one of the purposes of the stress tests is to promote confidence . . . whether that confidence is justified or not presumably being a different matter.

Revealing a "bad" result would also undermine confidence in the central bank itself, because it would raise awkward questions about its own competence. How could the banking system be so weak after all the resources devoted to rebuilding it, who should be held to account, and so forth? Of course, everyone understands that the central bank has a huge incentive to promote the message that the system is sound, and so the tests lack credibility because only a

reassuring answer is ever to be allowed. And, indeed, by curious coincidence, central bank stress tests *always* report that the banking system as a whole is sound, even if they sometimes report that individual banks are not.

Despite the fact that the principal purpose of central bank stress tests is to make the banking system more stable, those tests *inevitably* increase systemic banking risk:

- Such tests embody a regulatory risk standard—an "approved" way to manage risks—and any such standard is inherently counterproductive. In the absence of any standardization, different banks will have different approaches to risk management and this diversity will help to stabilize the system, with some buying in a crisis when others sell. But if all banks are pressured to standardize their risk management, then they will all act in much the same way and will all attempt to sell in a crisis. However, while an individual bank can sell in a crisis, the entire lot cannot. All assets have to be held by someone, and a collective attempt to sell only aggravates the crisis by exacerbating price instability. Moreover, *any* regulatory standard will have this destabilizing effect even if that standard is a good one when applied to any individual bank on its own.
- In practice, however, any such standard must inevitably be flawed—if only because no central bank has the incentives or information to produce a perfect standard—and in such cases the whole banking system will then be exposed to weaknesses in the "approved" risk management standard, and especially to the weaknesses in the "approved" models. They will then end up with the same flawed models with the same risk blind spots, and the entire system will be exposed to the dangers that the "approved" models fail to detect.

A Record of Repeated Failures

It is also instructive—indeed, grimly entertaining—to examine the track record of regulatory stress testing to see how this methodology actually works in practice.

The Freddie and Fannie Stress Tests

Let's begin with the first modern stress tests: the Fannie and Freddie tests. Their origins go back to the early 1990s, when there

was concern over the solvency of these government-sponsored enterprises (GSEs). There were proposals to increase their capital requirements, but Fannie managed to head off such pressure by means of an audacious coup. It commissioned former Fed Chairman Paul Volcker to examine the matter, and he concluded that Fannie was safe.

Fannie's chief executive could then claim that his business was safer than banking: "There are no unpleasant surprises because of the nature of our business. We don't have any see-through buildings, any Third World countries or any strip shopping malls. We just have those mortgages" (Hagerty 2013).

It then took nearly a decade for the rocket scientists to come up with model-based capital requirements that were not much higher than zero, and this when the GSEs were loading up on subprime mortgages, then known as "affordable housing." Fortunately, this was not a problem: The models said the subprime market was actually quite safe because modelers did not allow for any possibility of a housing downturn and the risks were diversified away.

As the details were being finalized, Fannie then scored another coup by commissioning a distinguished team of economists led by Joseph Stiglitz to carry out its own stress tests. The Stiglitz team came back with the reassuring conclusion that even under a decade-long "nuclear winter" scenario, the probability of Fannie or Freddie failing was essentially zero (Stiglitz, Orzag, and Orzag 2002).

The GSEs then went on a massive binge and effectively failed six years later when the government took them into conservatorships to avert impending collapse. The tests had been a spectacular failure.

So what went wrong? Well, part of the problem was that the stress-based capital requirements were way too low, but part of the problem was that the new system allowed the GSEs to game the system by loading up on risks that the models did not adequately capture. The GSEs' management teams were also working with contracts that encouraged excessive risk-taking so the outcome should be no surprise. They were also gaming the GSEs' government-sponsored status: They would tell Congress not to worry because the government was not on the hook, then tell Wall Street not to worry because the government *was* on the hook. There was also all the political meddling as well (see Morgenson and Rosner 2012). Or, to quote the book *Alchemists of Loss*, Fannie and Freddie

> leveraged more than would have been possible without the government's quasi-guarantee, used taxpayers' money to lobby like crazy to ensure they were not properly regulated and collapsed thankfully into the arms of the taxpayer as soon as the consequences of their own ineptitude became clear. It is indeed astonishing to consider how they managed to turn the soundest product in financial markets, the home mortgage, into a speculative casino, causing collateral damage of many times their own losses [Dowd and Hutchinson 2010: 190].

And all of which was missed by the stress tests.

The Fed's Stress Tests

The next important player to enter the stress-testing game was the Federal Reserve. Its first stress tests were carried out under the relatively light Supervisory Capital Assessment Program in 2009, followed by the CCAR in 2011, which has since become an annual event.

The CCAR is a highly aggressive program in which banks are required to prove the adequacy of their models relative to the Fed's models. Each CCAR has been more extensive and demanding than the previous one. Then in 2013, the CCAR was supplemented by additional stress tests mandated by Dodd-Frank, and in 2014 U.S. banks were subject to even more stress tests under Basel III.

Critics pointed out that the Fed's tests were reliant on the Fed's scenarios that were not particularly stressful, and were conditioned by political factors such as the Fed's reluctance to face up to the problems posed by the big zombie banks and the still unresolved problems in real estate markets. The tests were also blind to risks credibly identified by outside observers, for example:

- The risks of a eurozone collapse were ignored until the 2012 CCAR, and the eurozone had nearly collapsed the year before.
- The CCAR still ignores the biggest risk of all—that created by enormous off-balance sheet activities.

When I was researching *Math Gone Mad*, I interviewed some of the senior managers of a major U.S. bank. They told me that much of its normal activity had to stop because of the need to feed the models demanded by the Fed, and both its management and its IT systems were overwhelmed by compliance issues. The bank was

forced to make huge investments in models and modelers it didn't need, and then had to take more risks to recoup the costs. It also had to call a halt to further acquisitions because it couldn't assess the regulatory risks in potential purchases. Its whole business model became warped by the models, right down to the level of individual loans. And the models themselves couldn't be challenged.

Banks have no choice but to manage to what they perceive the Fed's models to be, otherwise they fail the tests. The result is the banks end up with much the same unreliable models, they then make much the same mistakes, and the U.S. banking system ends up with much greater systemic risk—a risk that none of the models pick up.

Over time, the tests also become routine and the results increasingly predictable. Stress testing then becomes a meaningless but very costly and very counterproductive compliance exercise. In fact, there is now a flourishing consultancy industry that specializes in how to pass the tests. The consultants are former Fed officials—the same ones who used to conduct the stress tests themselves. The very process of repeated stress testing over time has made the tests themselves futile. Put all this together, and you have lots of jobs for risk modelers, but the net effect is a growing systemic risk that the models cannot see. But let's not be too hard on the Fed—when it comes to screwing up stress testing, the Fed is only an amateur. Let's look at other countries.[3]

The Icelandic Stress Tests

Consider Iceland. In 2004, the three largest Icelandic banks had assets equal to 100 percent of GDP. They then embarked on a massive expansion and by the end of 2007 their assets were 900 percent of GDP—a world record. credit default swap (CDS) spreads were now suggesting that there might be a problem.

Fortunately, an IMF stress test in August 2008 suggested there was nothing to worry about—the system was resilient. So did a bunch of other regulatory stress tests. The banking system then collapsed unexpectedly in October: The stress tests had missed the imminent collapse of the entire Icelandic banking system!

[3]The discussion of the U.K. and European stress tests is drawn mainly from Dowd (2015).

The U.K. Stress Tests

Then there are the new kids on the stress-testing block, the Brits. The first U.K. stress test was carried out last year. The message was that the banking system was sound.

The exercise was based on a single scenario, and a mild one at that: GDP growth falls to −3.2 percent before bouncing back, inflation rises to peak at 6.5 percent, long-term gilts peak just below 6 percent, and unemployment hits 12 percent. This is not particularly stressful by historical standards, and also pales in comparison to the output falls and unemployment rates in parts of the eurozone. The impact of this stress scenario is also very mild: The average CET1/RWA ratio drops a little from 10 to 7.3 percent and there is a fairly small drop in bank profitability.

The Bank of England also uses a very low "pass" standard—a 4.5 percent minimum ratio of CET1/RWA. This is lower than the 5.5 percent ratio that the European Central Bank used in its widely discredited 2014 stress tests, is lower than the 7 percent ratio already in force in the United Kingdom in 2014, and is well below the minimum capital requirements coming through under Basel III. Had the Bank carried out a test using these latter Basel minima, however, the U.K. banking system would have failed: same exercise, higher safety standard, opposite result.

The Bank also failed to carry out any tests based on leverage—the ratio of capital to total unweighted assets—which offers a much less gameable measure of a bank's financial health. Even an undemanding test based on the Bank's required minimum leverage ratio of 3 percent would have revealed how weak the U.K. banking system really was, and most of the banks would have failed.

The Bank's failure to test against its own minima hardly inspires confidence. But then again, neither does the alternative. One might say that the Bank is damned because it didn't, and would have been damned if it did.

A 3 percent test is the weakest of leverage ratio tests. It is lower than the 4 percent minimum that the Fed now uses for the CCAR and a fraction of the 15 percent–plus minimum that many experts recommend (see, e.g., Admati et al. 2010 and Admati and Hellwig 2013: part III). The 3 percent minimum is at least five times larger than the leverage test that the Bank failed to conduct—or, at least, to report. Had the Bank based its stress tests on this measure, December's

comforting financial headlines would have been very different indeed. By this standard, the whole U.K. banking system would not so much be underwater as stuck in Davy Jones' locker at the bottom of the ocean. We can therefore safely dismiss the U.K. stress tests.

The European Stress Tests

The European stress tests are even worse. The first of these was carried out by the Committee of European Banking Supervisors in 2009. The outcome was uneventful: 22 large banks all passed. European regulators were quick to hail the "resilience" of the European banking system. Critics simply said that the stress test was too weak.

The fun starts with the second stress test conducted in 2010. This test covered the 91 largest European banks and only seven failed to meet the 6 percent minimum. Their combined shortfall was only €3.5bn—about 0.15 percent of GDP. This figure was a fraction of the estimates of independent analysts, and the stress test largely ignored the biggest risk of all—the risk of sovereign defaults. This risk was ignored, it turned out, because the EU was "committed" to ensuring that such defaults never happened—a classic case of policy make-believe undermining the exercise from its inception. Any doubts about its credibility were dispelled four months later when the Irish banking system collapsed. The Irish banks had all passed the stress test.

The 2011 stress tests were then carried out by the new European Banking Authority (EBA). The get tough EBA promised that its stress tests really would be credible and it would not repeat the mistakes of the earlier fiascos. There was now a much greater awareness of the sovereign debt problem and the EBA needed to prove itself. So what did it do? It came out with an aggregate shortfall of €2.5bn, even *less* than the widely discredited estimate from the year before. Any doubts about the credibility of that exercise were then dispelled three months later when the giant bank Dexia failed. Dexia had aced the test.

In the meantime, the EBA frantically revised its numbers. Its corrected estimate turned out to be over 45 times larger than the original—and even this was much less than independent estimates. Then the big Spanish bank Bankia failed; Bankia had also passed. Then the icing on the cake—the entire Cypriot banking system

collapsed in early 2013; the big Cypriot banks had also passed the test. None of the key agencies monitoring Cyprus—the EU, EBA, IMF, and BIS—even had Cyprus on any kind of watch list. So now the stress testers had a hat trick. Three national banking systems had failed after being signed off as sound.

The next major EU stress tests were conducted by the ECB in 2014 as part of its new mandate as Europe's super-regulator. A key driver behind the establishment of the eurozone banking union and the Single Supervisory Mechanism to govern it was the argument that national regulators were prone to capture and therefore an independent and more demanding regulator was required—namely, the ECB. The ECB promised that *its* stress tests really would be credible and it would not repeat the mistakes of the earlier stress-test fiascos. The ECB stress test was also to be buttressed by an asset quality review (AQR) to provide assurance that the new stress tests would be based on sound data given the glaring data problems that had plagued earlier stress tests. The new tests were also to have a stronger capital standard, an 8 percent CET1/RWA hurdle ratio—the standard minimum of 4.5 percent, plus a 2.5 percent CCB, plus a 1 percent G-SII requirement. Unfortunately, the 8 percent ratio soon attracted a lot of negative lobbying from interested parties—the banks and their national supervisors, who had been captured by them—and the hurdle ratio was eventually knocked down to 5.5 percent.

Twenty-five banks then failed the stress test with a combined shortfall of €25 billion. None of the biggest banks failed, and the banks that did fail were concentrated mainly in the southern fringe. For its part, the asset quality review produced asset quality adjustments of an additional €48 billion. The severity of the stress is apparent when one considers that the combined shortfall plus quality adjustment amounted to only about 0.3 percent of total bank assets, a number small enough to be rounding error. One wonders why they bothered. In any case, a chorus of experts dismissed the results on publication.

One problem was the adverse scenario, which was very mild. This scenario assumed that inflation would drop to 1 percent in 2014. But by the time the results were released inflation had fallen well below this level and much of the eurozone was already in deflation. When challenged about this at the press conference, ECB Vice President

Constancio's response makes satire redundant: "The scenario of deflation is not there because . . . we don't consider that deflation is going to happen," he said.

Independent experts estimated shortfalls nearly 30 times larger than the ECB's estimates (Acharya and Steffen 2014a, 2014b, Vestergaard and Retana 2013). These assessments are superior because they use standardized, easily replicable low-cost approaches and are credible because they are independent of the political influences and regulatory capture that compromise central bank stress tests.

Independent studies also suggest that the biggest risks are in the French and German banks, directly contrary to the ECB party line that the core eurozone banks are sound and that any problems lie primarily around the southern fringe.

The main reason for the discrepancy between their results and the ECB's is that they used leverage ratios that *revealed* the risks rather than RWA ratios that *hid* them. The implication is that these banks only appear strong because of their superior expertise in gaming the risk weights.

Consider the big French and German banks, Credit Agricole, BNP Paribas, SocGen and Deutsche.[4] Each of these

- easily passed the ECB stress test,
- would easily have failed an undemanding 3 percent leverage ratio test,
- would produce enormous shortfalls under a severe (7 percent) leverage ratio test, and
- had very low ratios of risk weighted to total assets.

In sum, these banks are more risky, but better at making their risks invisible to the ECB test.

[4]Deutsche is a particular problem child. It is arguably the largest, most systematically important bank in the world, with a total balance sheet about the same size as world GDP. With deposits only 1 percent of its balance sheet, it is a gigantic hedge fund with a comparatively small bank attached. FDIC Vice Chairman Thomas Hoenig has been warning about it for years; two years ago, he noted: "It's horrible, I mean they're horribly undercapitalized. They have no margin of error." Recent investigations by the New York Fed also indicate a litany of serious problems in its U.S. arm, indicative of the firm worldwide. These include shoddy reporting, inadequate auditing and oversight, and weak technology systems, amounting to a "systemic breakdown" in controls (see Enrich, Strasbourg, and Henning 2014). Deutsche is, in short, a disaster waiting to happen.

It would therefore appear that that the stress tests had been driven and hence compromised by the desire not to offend powerful governments—especially Germany and France—who also had their own reasons to want test results suggesting that the problems lay on the fringes of the eurozone, and not right at its heart. That same message would have also suited the empire-builders at the ECB to reinforce the case for giving them yet more power. In any case, it would have suited no one for the ECB to suggest that some of Europe's too-big-to-fail banks were, well, on the verge of failure, as that would have put the spotlight on them to come up with a solution to this most vexing of problems. And so the

> suspicion lingers that undertaking the comprehensive assessment on the basis of risk-weighted assets and an only mildly adverse stress scenario were not "mistakes," after all. More likely, it reflects substantial political pressures. It would have required courage and genuine independence for the ECB to identify several German and French banks as severely undercapitalized just days before it assumes bank supervisory responsibilities for all major eurozone banks [Vestergaard 2014].

In short, the ECB had been captured and its stress tests were no more credible than its predecessors' had been.

Conclusion

The inescapable conclusion is that the methodology of regulatory stress testing is based on foundations that are indefensible and even risible. These include:

- Their dependence on discredited models of financial risk;
- Their reliance on a single scenario, against all good practice and common sense;
- Their tendency to rely on unstressful stress scenarios and very low pass marks;
- Their intrinsic lack of credibility because of regulatory capture, central banks' own dismal forecasting records, and the politics that underlie and fatally undermine regulatory regulatory stress exercises; and
- Their reliance on "approved" models and risk management practices that increase systemic risk—and do so in a way that the models themselves cannot see.

If this does not persuade, consider their track record, which includes, among other embarrassing disasters, the hideously costly Fannie and Freddie stress-test fiasco and three national banking systems signed off as sound by regulatory stress tests, which then collapsed unexpectedly not long afterwards. In fact, with one exception, I am not aware of a single case in which a regulatory stress test correctly identified *in advance* a subsequent buildup of banking stress, which then allowed the relevant authorities the chance to counter it. The exception? The Northern Rock "war games" in the United Kingdom in 2005. This exercise identified a scenario very close to that which was to lead to the Northern Rock run in 2007—the first English bank run since 1866—but then the British authorities did nothing about it and were caught completely unprepared when their own scenario came to pass.

In not a single case did a regulatory stress exercise ever lead to any demonstrated benefit to the banking system. On the other hand, there are many cases where such exercises provided false comfort, lulling those involved to sleep in the face of imminent danger. Ironically, they generally did so when more conventional indicators—such as CDS or yield spreads, and leverage and other ratios—were clearly indicating red. But the stress tests said not to worry. So they didn't and then disaster struck.

Stress tests operate like a cancer detection procedure that can't identify most cancers, or a radar system that cannot see many of the hazards out there. We would never send out a ship or plane that relied on a radar that didn't work; we really shouldn't allow central banks or bank regulators to do the same with our banking systems either.

References

Acharya V. V., and Steffen, S. (2014a) "Falling Short of Expectation: Stress Testing the Eurozone Banking System." *Vox* (21 January).

_____ (2014b) "Benchmarking the European Central Bank's Asset Quality Review and Stress Test: A Tale of Two Leverage Ratios." *Vox* (21 November).

Admati, A.; Allen, F.; Brealey, R.; Brennan, M.; Boot, A.; Brunnermeier, M.; Cochrane, J.; De Marzo, P.; Fama, E.; Fishman, M.; Goodhart, C.; Hellwig, M.; Leland, H.; Myers, S.; Pfleiderer, P.;

Rochet, J. C.; Ross, S.; Sharpe, W.; Spatt, C.; and Thakor, A. (2010) "Healthy Banking System Is the Goal, not Profitable Banks." *Financial Times* (9 November).

Admati, A., and Hellwig, M. (2013) *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It*. Princeton, N.J.: Princeton University Press.

Dowd, K. (2014) *Math Gone Mad: Regulatory Risk Modeling by the Federal Reserve.* Cato Institute Policy Analysis No. 754. Washington: Cato Institute.

_____ (2015) *No Stress: The Flaws in the Bank of England's Stress Testing Programme.* London: Adam Smith Institute.

Dowd, K.; Cotter, J.; Humphrey, C.; and Woods, M. (2008) "How Unlucky is 25-Sigma?" *Journal of Portfolio Management* 34 (4): 76–80.

Dowd, K., and Hutchinson, M. (2010) *Alchemists of Loss: How Modern Finance and Government Intervention Crashed the Financial System*. New York: Wiley.

Enrich, D.; Strasbourg, J.; and Henning, E. (2014) "Deutsche Bank Suffers from a Litany of Reporting Problems, Regulators Said." *Wall Street Journal* (22 July).

Finger, C. (2008) "Epilogue: Fishing for Complements." In D. Rösch and H. Scheule (eds.) *Stress Testing for Financial Institutions: Applications, Regulations and Techniques*, 443–52. London: Riskbooks.

Hagerty, J. R. (2013) "How Paul Volcker Misread Fannie Mae." *Wall Street Journal* (7 September).

Haldane, A. C. (2013) "Containing Discretion in Bank Regulation." Speech given at the Atlanta Fed conference, "Maintaining Financial Stability: Holding a Tiger by the Tail(s)" (9 April).

Kerr, G. (2010) "How to Destroy the British Banking System: Regulatory Arbitrage via 'Pig on Pork' Derivatives." (21 January 21): www.cobdencentre.org/2010/01/how-to-destroy-the-british-banking-system.

_____ (2011) *The Law of Opposites: Illusory Profits in the Financial Sector*. London: Adam Smith Institute.

Larsen, P. T. (2007) "Goldman Pays the Price of Being Big." *Financial Times* (13 August).

Morgenson, G., and Rosner, J. (2012) *Reckless Endangerment: How Outsized Ambition, Greed, and Corruption Created the Worst Financial Crisis of Our Time*. New York: St. Martin's Press.

Stiglitz, J.; Orzag, J. M.; and Orzag, P. R. (2002) "Implications of the New Fannie Mae and Freddie Mac Risk-based Capital Standard." *Fannie Mae Papers* 1 (2): 1–10.

Vestergaard, J. (2014) "European Banking Misery: Pretending Rather than Mending Does No Favours to Lending." *GEG Watch* (4 November).

Vestergaard, J., and Retana, M. (2013) "Behind Smoke and Mirrors: On the Alleged Recapitalization of Europe's Banks." Danish Institute for International Studies *Report 2013:10.* Copenhagen: Danish Institute for International Studies.